Automated, on-board terrain analysis for precision landings

Zia-ur Rahman[†], Daniel J. Jobson[‡] Glenn A. Woodell[‡], Glenn D. Hines[‡],

[†]College of William & Mary, Department of Applied Science, Williamsburg, VA 23187,
[‡]NASA Langley Research Center, Hampton, Virginia 23681.

ABSTRACT

Advances in space robotics technology hinge to a large extent upon the development and deployment of sophisticated new vision-based methods for automated in-space mission operations and scientific survey. To this end, we have developed a new concept for automated terrain analysis that is based upon a generic image enhancement platform—multi-scale retinex (MSR) and visual servo (VS) processing. This pre-conditioning with the MSR and the VS produces a "canonical" visual representation that is largely independent of lighting variations, and exposure errors. Enhanced imagery is then processed with a biologically inspired two-channel edge detection process, followed by a smoothness based criteria for image segmentation. Landing sites can be automatically determined by examining the results of the smoothness-based segmentation which shows those areas in the image that surpass a minimum degree of smoothness. Though the MSR has proven to be a very strong enhancement engine, the other elements of the approach—the VS, terrain map generation, and smoothness-based segmentation—are in early stages of development. Experimental results on data from the Mars Global Surveyor show that the imagery can be processed to automatically obtain smooth landing sites. In this paper, we describe the method used to obtain these landing sites, and also examine the smoothness criteria in terms of the imager and scene characteristics. Several examples of applying this method to simulated and real imagery are shown.

Keywords: multi-scale Retinex, visual servo, segmentation, terrain classification, smoothness coefficient

1. INTRODUCTION

A fundamental concern in the development of resilient, vision-based, automation technology is the impact of wide-ranging extraneous lighting and exposure variations on the acquired imagery. This concern can be considerably ameliorated by the application of the multi-scale retinex (MSR)[1–5] algorithm. The MSR is a non-linear, context-dependent enhancement algorithm that provides color-constancy, dynamic range compression and sharpening:

$$R_i^M(x_1, x_2) = \sum_{k=1}^{K} W_k R_i(x_1, x_2; k), \quad i = 1, \ldots, \mathcal{B} \tag{1}$$

$$R_i(x_1, x_2; k) = \log\left(I_i(x_1, x_2)\right) - \log\left(I_i(x_1, x_2) * G_k(x_1, x_2)\right) \tag{2}$$

where R^M is the MSR output, K is the number of scales, W_k is the weight of each single-scale retinex (SSR) output, R, I is the input image, $\mathcal{B}$ is the number of spectral bands in the image, and

$$G_k(x_1, x_2) = \kappa_k G(x_1, x_2; k) \tag{3}$$

$$G(x_1, x_2; k) = \exp\left(-(x_1^2 + x_2^2)/\sigma_k\right), \tag{4}$$

$$\kappa_k = \left[\sum_{x_1, x_2} G(x_1, x_2; k)\right]^{-1}. \tag{5}$$

The σ_k are scale parameters that control the performance of the SSR. As can be seen from the examples in Figure 1, small σ_k lead to SSR outputs that contain all the fine features in the image at the cost of color, and

Contact: ZR: `zrahman@as.wm.edu`; DJJ: `daniel.j.jobson@nasa.gov`; GAW: `glenn.a.woodell@nasa.gov`; GDH: `glenn.d.hines@nasa.gov`;
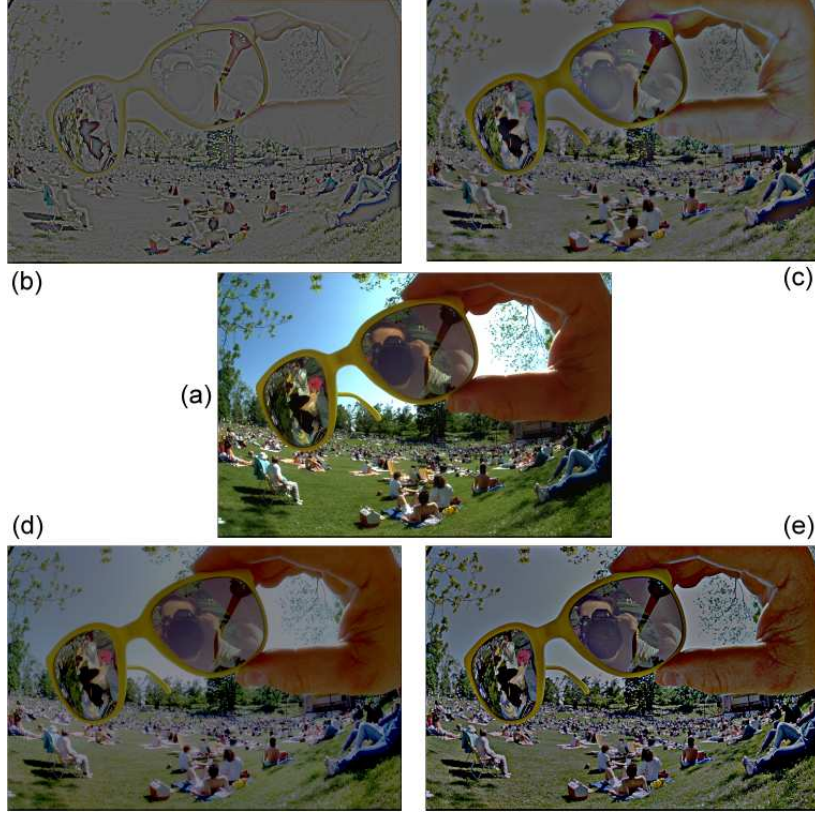
Figure 1. (a) The original input (b) Narrow surround $\sigma = 5$ (c) Medium surround $\sigma = 20$ (d) Wide surround $\sigma = 240$ (e) MSR output with $W_k = 1/3, k = 1, 2, 3$. The narrow-surround acts as a high-pass filter, capturing the fine detail in the image but at a severe loss of tonal information. The wide-surround captures the fine tonal information but at the cost of fine detail. The medium-surround captures both dynamic range and tonal information. The MSR is the average of the three renditions.

large σ_k lead to outputs that contain color information, but not fine detail. There is an obvious problem with the MSR output shown in Figure 1(e): where the input image contains areas of significant extent that have constant or slowly varying intensity values, the MSR output tends to middle gray. However, this phenomenon is not a concern for terrain analysis where the color information is not being used. The inherent color-constancy and illuminant independence of the MSR[1,3] is shown in Figure 2.

In order to extract consistent scene structure from any image under widely varying scene and sensor conditions, one has to think in terms of transforming the image into a "canonical" representation that effectively eliminates such undesirable variability. The MSR has proven to be a powerful tool for doing just this. Because of its dynamic range compression and illumination independence properties, the MSR provides consistent rendering for imagery from highly diverse scene and sensor conditions. However, the MSR was designed to handle moderate to very wide dynamic range images: It was not designed to deal with the case of narrow dynamic range images encountered in turbid imaging conditions such as fog, smoke, and haze, dim lighting conditions, or significant under- or over-exposures. To expand the performance envelope of the MSR, we have developed a "smart" framework of visual quality measurements and enhancement controls that we call the Visual Servo[6–8] (VS). The VS assesses the quality of the image in terms of brightness, contrast and sharpness, and controls the strength of the MSR enhancement. This combination of servo with retinex—SRVR—is, therefore, an appropriate platform for vision-based automation under wide-ranging extraneous variability—see Figures 3 and 4.

Figure 3 shows examples of imagery acquired from airborne platforms. The original images are dark, almost mono-chromatic, and blurred. The SRVR processed images bring out the details and the color in the imagery
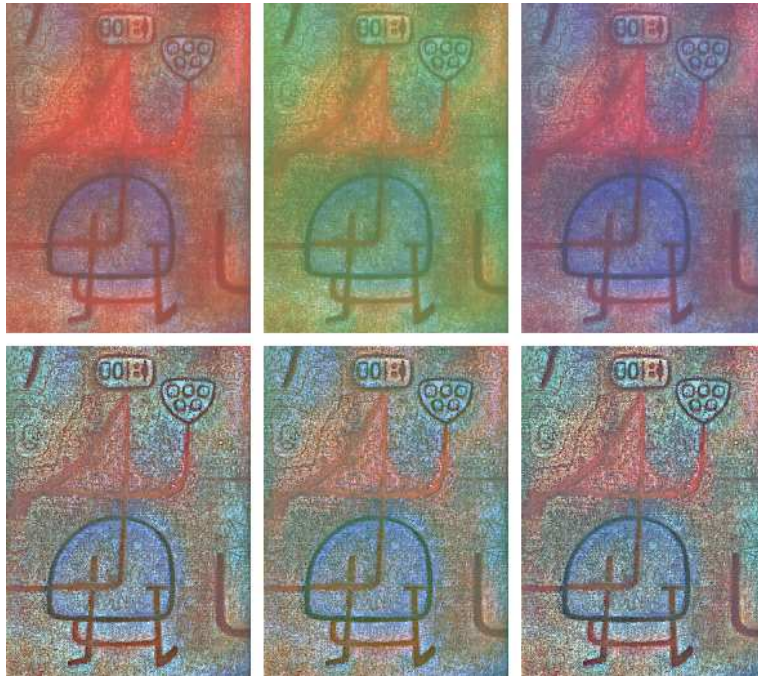
Figure 2. The image shows a painting by the Swiss artist Paul Klee, where the effect of changing the illuminant was simulated by red, blue, and green shifting the original image (top row). The bottom row shows the MSR output for each case. Note that the MSR outputs are *almost* color constant—much like the human visual system.
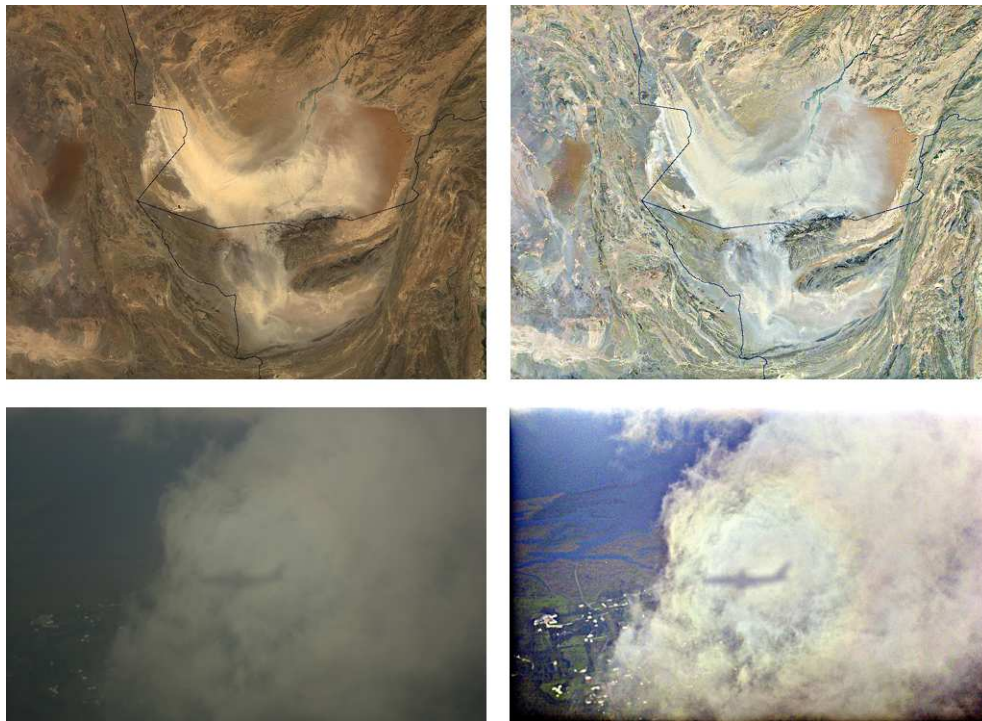
Figure 3: Examples of automatic SRVR image enhancement for wide ranging lighting conditions.

Figure 4. Demonstration of exposure error compensation for declining in light levels. While camera aperture, shutter speed, and ISO setting were constant over this sequence, camera position was not.

and show details in areas where no detail was visible in the original data. Figure 4 shows an example of the capability of the SRVR to correct for simple exposure error. It presents a series of images of the same scene acquired from twilight to darkness. The SRVR enhancements produce a very robust, consistent visual rendition that varies only in the increasing sensor noise levels accompanying the dropping light levels. Because the SRVR employs strong degrees of non-linear spatial processing, it handles "within" image variability with a high degree of agility and succeeds in forming visual renditions that are dominated by the scene reflectance and topography content.

While terrestrial imaging is rich in both reflectance and topographical variations, lunar and planetary imaging is overwhelmingly dominated by topographical variations only. Further, any reflectance variations that do occur—for example, rocks vs. dust on the surface of Mars—are usually associated with a topographic relief. Therefore, we can reasonably assume that signal variations at the scale of orbital image surveys of lunar and planetary surfaces, equate topographic relief with sharpness and degree of local image contrast. Larger and steeper signal differences are associated with rugged terrain; while signal variations of lower contrast and slope are due to more rolling terrain. Smooth terrain will largely be devoid of signal variations. This set of assumptions is the basis for the automatic terrain classification scheme described in this paper. Obviously, the occurrence of a sharp reflectance boundary—deep shadow—on smooth terrain, will cause it to be misclassified as rough terrain. This misclassification will be considerably ameliorated by the application of the MSR which will bring out the details in the dark regions. In any case, such errors are likely to be rare and can be corrected with subsequent verification analysis by the human specialists. In the case of such an occurrence during a landing approach, the system will (rightly) "err on the side of caution" and steer the landing away from such terrain.

In the remainder of this paper, we will describe the automatic terrain analysis methodology (Figure 5) and initial results from its application to diverse Mars terrain using image data from both the Mars Global Surveyor[9] and the more recent Mars Express[10] orbiting imagers. We will illustrate the noise evasion properties of image-adaptive edge detection and the ability of such adaptiveness to maintain relative topographic variations in the presence of high relief that might otherwise saturate the terrain scale. Hundreds of diverse terrain images from Mars Global Surveyor and several from the Mars Express images have been tested and inspire confidence that this approach and these methods are indeed capable of achieving the desired resilient, vision-based, automation for terrain analysis of the Moon and planets. Obviously, gaseous planets such as Jupiter and Saturn are not going to involve surface survey by imaging, but many of their moons will. These methods, however, are not going to be relevant to the optically thick atmospheres of Venus, Titan, and the like, unless they can be found useful in the radar imaging used for those cases. This issue is outside the scope of our present efforts.

2. AUTOMATED TERRAIN ANALYSIS

The process for automated terrain analysis is shown in Figure 5. All input imagery is first SRVR processed. The primary purpose of SRVR processing, in relation to scene structure extraction, is to produce a visual representation of the scene that is (approximately) "canonical," i.e., largely independent of within-frame and "frame-to-frame" lighting variations, and exposure errors (Figure 4). This provides a resilient computational substrate that dramatically improves the consistency of edge detection and structural scene description. An adaptive edge detection process is then applied to the SRVR enhanced image (Section 2.1). The edge-image is then segmented using a smoothness criteria (Section 2.3) to detect the presence of sufficiently large areas that can be used at potential landing sites.
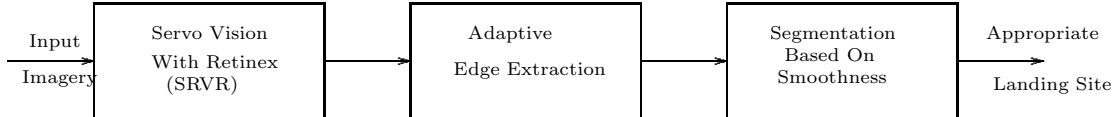

Figure 5: Overview of Automated Terrain Analysis.

2.1. Edge-extraction for terrain analysis

We use a two-scale edge detection process that mimics human visual perception of edges and is inspired by the parvo-cellular, P-channel, and magno-cellular, M-channel, physiological subsystems of natural vision. This is both an anatomical distinction[11] and a physiological distinction.[12, 13] In primate vision, the P-channel is associated with high acuity and coarse contrast sensitivity, while the M-channel is associated with coarse spatial resolution and high contrast sensitivity. This arrangement of the channels appears to eliminate all except large noise spikes from the P-channel, and provides some noise smoothing benefits due to spatial averaging in the M-channel. Edge events detected in the P-channel are associated with sharp, high-relief terrain that represents significant topographic obstacles to safe landing. Events detected in the M-channel, but not the P-channel, are indicative of "rolling" terrain that may or may not be suitable for landing, depending upon specific mission parameters. For example, a robotic craft can probably accept rougher landing than is is generally acceptable for manned operation. Finally, regions that have few to no P events and minimum number of M events can be classified as "smooth" and be added to the list of candidate landing sites.

The specific edge detection operator in this paper is a center-surround "difference-of-Gaussians" (DOG), $\mathcal{G}$. In the frequency-domain,

$$\hat{\mathcal{G}}(\omega_1, \omega_2; \rho_c; \zeta) = \exp\left[-(\omega_1^2 + \omega_2^2)/\rho_c^2\right] - \zeta \exp\left[-2.56(\omega_1^2 + \omega_2^2)/\rho_c^2\right] \tag{6}$$

where $\hat{\mathcal{G}}(\omega_1, \omega_2)$ is the Fourier transform of $\mathcal{G}$ at frequencies (ω_1, ω_2), ρ_c controls the extent of $\mathcal{G}$, and ζ controls the weighting of the difference operation. The first term on the right-hand-side of Equation 6 can be thought of as the optical transfer function (OTF), $\hat{\tau}$, of the imaging system, where ρ_c controls the amount of blur the system produces. Small values of ρ_c lead to more blur, and vice versa. Figure 6 shows $\mathcal{G}$ for $\rho_c = 0.30$, and $\rho_c = 0.20$ as a function of the parameter ζ. When $\zeta = 1$, $\mathcal{G}$ is the classic "Mexican hat" spatial response function which, when convolved with the input image, locates edges at zero-crossing points in the output image. This response is similar to that of the P-channel for $\rho_c = 0.30$. When $\zeta = 0.5$ and $\rho_c = 0.20$ (Figure 6(b)), then $\mathcal{G}$ allows a blurred version of the original image to pass through, and the response is then similar to the M-channel processing. The spatial scales of the center and surround determine the spatial acuity of the edge detection process, while an edge detection threshold determines contrast sensitivity.

2.2. Adaptive thresholding for edge-detection

In order to determine whether the acquired imagery is of an area that is safe to land on, a two-part process is invoked. In the first part, adaptive edge detection is invoked if the initial global edge-count exceeds a pre-set value for either the P or M channel, or for both. This is indicative of a scene that has a high degree of topographic
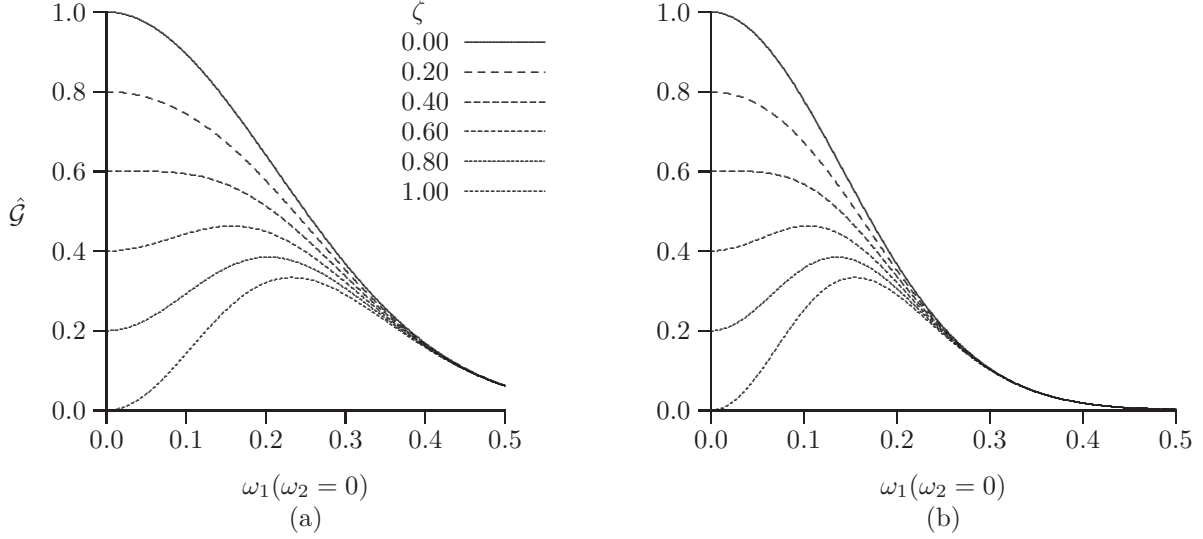
Figure 6: The Difference of Gaussian operator: (a) $\rho_c = 0.30$; (b) $\rho_c = 0.20$.

relief, or significant noise. In this adaptive process, edge detection thresholds[*] are allowed to drift until the total edge counts fall below their pre-set global count criteria. This is controlled to a limited range in order to avoid rugged terrain being mis-characterized as "rolling" or "smooth." In this manner, adaptiveness over a controlled range preserves relatively smoother terrain with lower edge counts than the more rugged terrain.

Likewise, the adaptive edge detection provides a degree of noise evasiveness. For images with moderate noise levels that create many spurious edge events unrelated to terrain features, the drifting upward of edge thresholds would reduce the number of "false positive events" while preserving significant terrain features. For the case of really high levels of noise, this strategy no longer works and the image data is dominated by noise-generated "false" edges. In this case, all such image regions will still be classified as "rugged" terrain and be edited out of potential landing site selections. This seems to be the best disposition since such noisy data means scientifically that the terrain is still unknown, and this overall scheme then is understood to "err on the side of caution" by keeping unknown terrain out of the landing site selection.

2.3. Smoothness based segmentation

Once the edge-image has been obtained, a smoothness condition is used to partition the image using a quad-tree segmentation algorithm.[14] The smoothness criterion is directly derivable from the edge representation of the image. The smoothness coefficient, $\mathcal{S}$ is given by:

$$\mathcal{S}^2 = \frac{1.0}{\sum_{\omega_1,\omega_2} \left| \hat{\mathcal{I}}(\omega_1, \omega_2) \hat{\mathcal{G}}(\omega_1, \omega_2; \zeta = 1.0; \rho_c = 0.3) \right|^2} \tag{7}$$

where, $\hat{\mathcal{I}}$ is the Fourier transform of the input image, and $\hat{\mathcal{G}}$ is given by Equation 6. The smoothness coefficient, $\mathcal{S}$, represents the reciprocal of the amount of energy in the high-pass filtered version of the input image, $\mathcal{I}$: the higher the amount of energy, the more the high frequency information. Since, the high-frequency information is directly correlated to the fine details in an image, the smoother an image the larger the $\mathcal{S}$.

An acquired image can be quantified in terms of its mean-spatial detail μ, the average distance between intensity changes (edges). Additionally, it is a function of the characteristics of the acquisition device such as

---

[*]An edge-detection threshold is defined by the signal difference above which an event is classified as an edge.
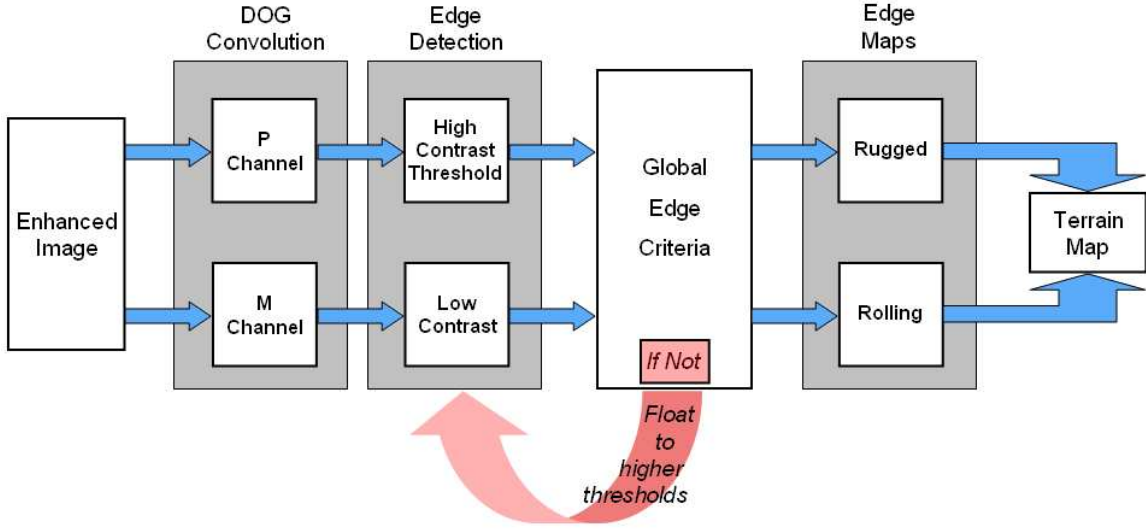
Figure 7: Processing for edge detection and classification.

the signal-to-noise ratio (SNR) and the optical transfer function $\hat{\tau}$. In order to quantify $\mathcal{S}$ in terms of the SNR and μ, we use simulated data that have a circularly symmetric power spectral-density, (PSD), Φ_L, given by,[15, 16]

$$\hat{\Phi}_L(\omega_1, \omega_2) = \frac{2\pi\mu^2\sigma_L^2}{\left[1 + \left(2\pi\mu(\omega_1^2 + \omega_2^2)\right)^2\right]^{3/2}}, \tag{8}$$

where σ_L is the energy that is received by the imager. This PSD has been shown to match that of natural scenes as recorded by a remote imager..[17] A set of images that have this Φ_L are shown in Figure 8 for a range of values of μ. Since the amount of detail in the image changes with the parameter μ, we can simulate situations where the terrain can be used for a safe landing, as well as situations where that is not possible. We can inject (white, Gaussian) noise into the imagery to simulate the noise in the camera, and also inject illumination changes across the images to simulate changes in lighting in the scene due to shadows cast by objects.

Figure 9 shows the variation in $\mathcal{S}$ as a function of μ for various SNRs. The "Ideal" line represents the noise-free case, and represents the upper bound of performance. The symbols represent the average value of 100 instantiations of random polygon images—like the ones shown in Figure 8—that have the PSD given by Equation 8, for a given μ and SNR. The continuous lines show a 3-point quadratic fit to the discrete data. As the SNR increases, the variation of $\mathcal{S}$ with μ becomes almost linear. Since the mean spatial detail in the image is a function of the height of the imager above the scene and the sampling lattice of the camera, the smoothness coefficient that determines whether the terrain is smooth or not is also a function of the height of the sensor above the scene. An examination of the simulated imagery shows that terrain where $\mu < 15$ is not suitable for landing. Using that as the baseline, we can determine for a given camera SNR, a minimum smoothness coefficient value that can be obtained from Figure 9 that must be surpassed for safe landing.

Figure 10 presents the same data that is shown in Figure 9 but as a function of the SNR for several different μ. Again, the "Ideal" line represents the noise-free case. For scenes with high spatial detail, i.e., small μ, the SNR does not have a large impact on $\mathcal{S}$. In fact, for $\mu = 1$ and $\mu = 5$, an imager with an SNR performance of 30 would be sufficient since a better imager would still not be able to distinguish between noise and signal detail. However, for smoother scenes, one would want to avoid the false alarms that a noisy imager would produce. Since we claim that areas with $\mu < 15$ are not safe for landing, we should design our imager to have an SNR performance of at least 100.

The smoothness coefficient is used to segment the image into regions that possess a minimum smoothness value. Experimental analysis shows that using $\mathcal{S} = 5$ is a reasonable choice for determining whether a region
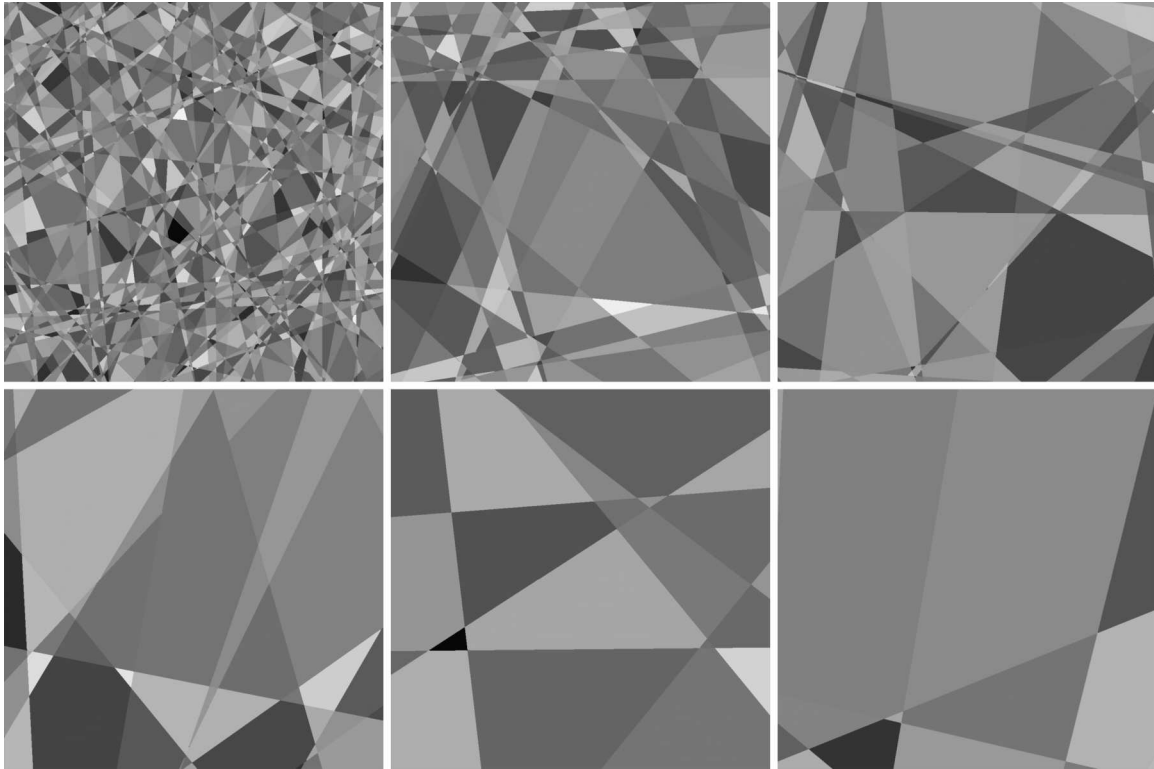
Figure 8. Top row: $\mu = 1, 3, 5$. Bottom row: $\mu = 9, 14, 20$. The detail in the images varies inversely with μ. The top row simulates the situation where there is no smooth terrain. The bottom row simulates situations where smooth terrain large enough for a safe landing can be found.

is smooth. This is in agreement with the data shown in Figure 10 where it can be seen that $\mathcal{S} > 5$ indicates a region with $\mu > 12$ for all SNRs except SNR $= 1$. Figure 11 shows the segmentation when the smoothness criteria $\mathcal{S}$ is used to segment the images in Figure 8.
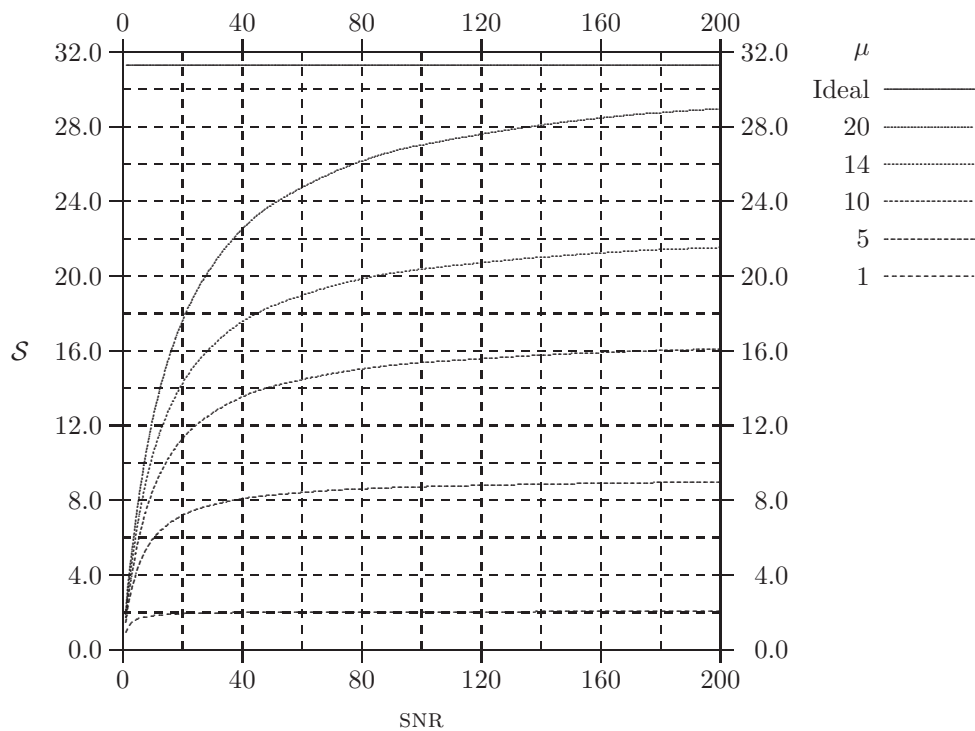
3. RESULTS

The automated terrain analysis methods have been applied to several hundred Mars Global Surveyor MOCS orbital images, and several of the orbital images of the current Mars Express HRSC. The results strongly support the hypothesis that the SRVR processing provides an excellent computational platform for resilient, consistent edge detection and scene topographic analysis, and that the automated terrain analysis method yields robust planetary topographic survey results. Figure 12 shows the original MGS image, the interim result of SRVR enhancement and the final result of terrain maps produced for highly diverse terrain types on Mars. A more extensive portfolio of results can be viewed at `http://dragon.larc.nasa.gov/retinex/marsterrain`.

Figure 13 shows the result of the segmentation operation on some images from the Mars Global Surveyor. The first-column shows the original image; the second-column shows the results of the smoothness-based segmentation of the original image; the third-column shows the SRVR enhancements; and, the fourth-column shows the segmentation of the SRVR enhanced imagery. It is interesting to note that in each of the three cases, segmentation of the original imagery shows that there are significant "smooth" areas to make landing feasible. However, this is not the result that is obtained from the segmentation of the SRVR enhanced imagery: regions that are classified as smooth in the original imagery have significant variations in the terrain that become evident only after enhancement. This is especially the case in the images shown in the third row where the poor imagery falsely suppresses all the variability of the terrain.
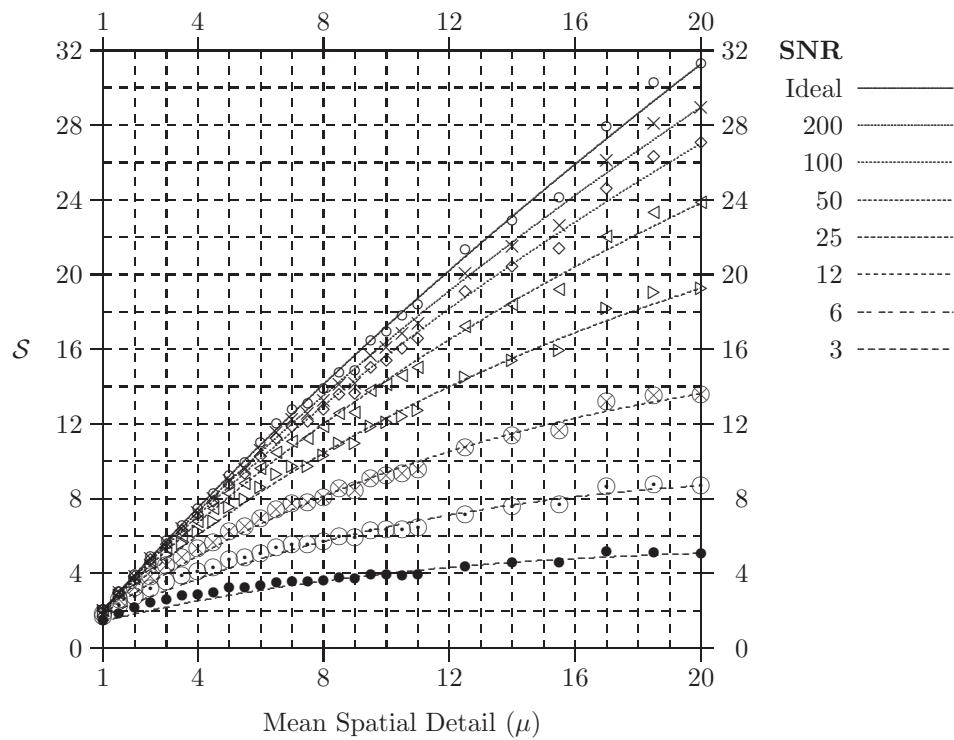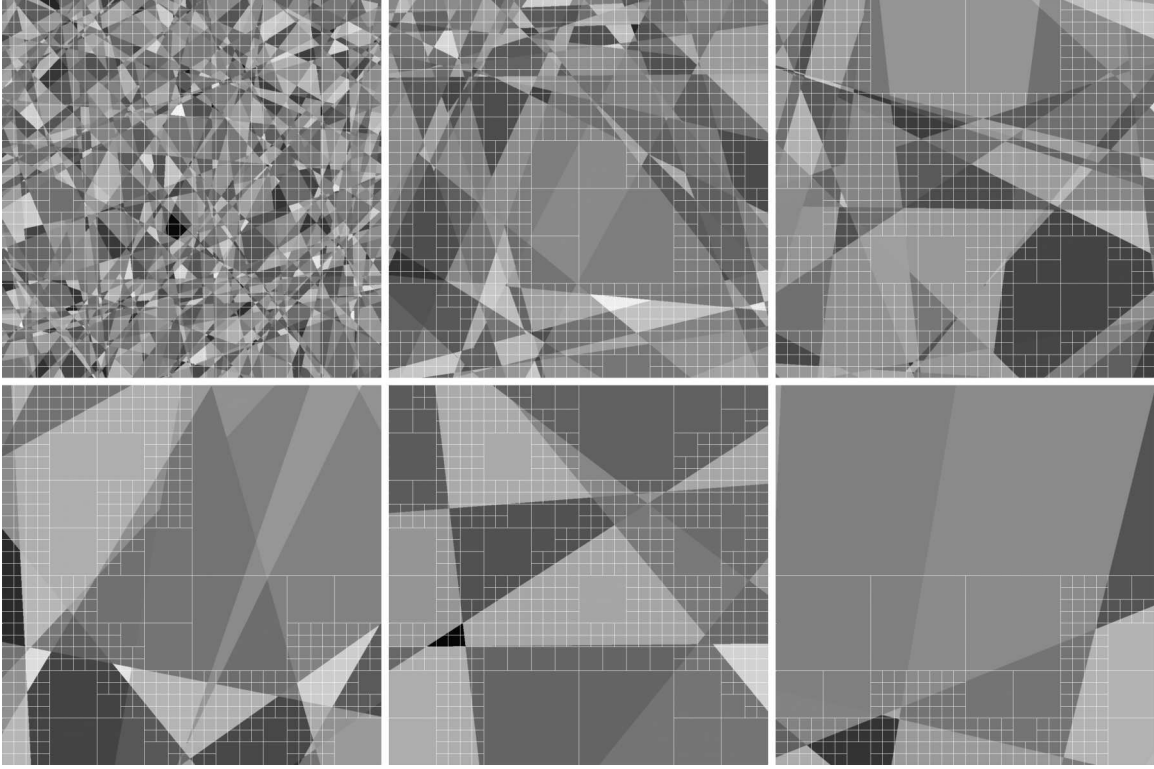
Figure 9: Smoothness as a function of mean spatial detail for changing SNR.

**Figure 10**: Smoothness as a function of SNR for changing mean spatial detail.

Figure 11. Segmentation of the scenes from Figure 8 for $\mathcal{S} = 5$ and $\mathcal{G}(\rho = 0.30; \zeta = 1.00)$. The segmentation closely follows the anticipated results with large "smooth" areas showing up only for large values of μ.

4. CONCLUSIONS

We have presented a scheme for automated planetary terrain analysis that uses image enhancement in combination with edge-analysis to determine if the selected terrain is actually smooth enough to land on. The approach builds on the MSR and the VS to transform imagery into a "canonical" visual representation that is largely independent of extraneous within-image and frame-to-frame lighting variations and exposure errors. Experiments show that this provides a very solid foundation for consistent edge extraction and hence, terrain characterization. We have tested this automated terrain analysis method extensively on Mars Global Surveyor and Mars Express image data for diverse terrain types. The results of these tests serve as a test-bed demonstration of the resiliency and utility of these methods for onboard planetary topographic survey from orbiting spacecraft.

Acknowledgments
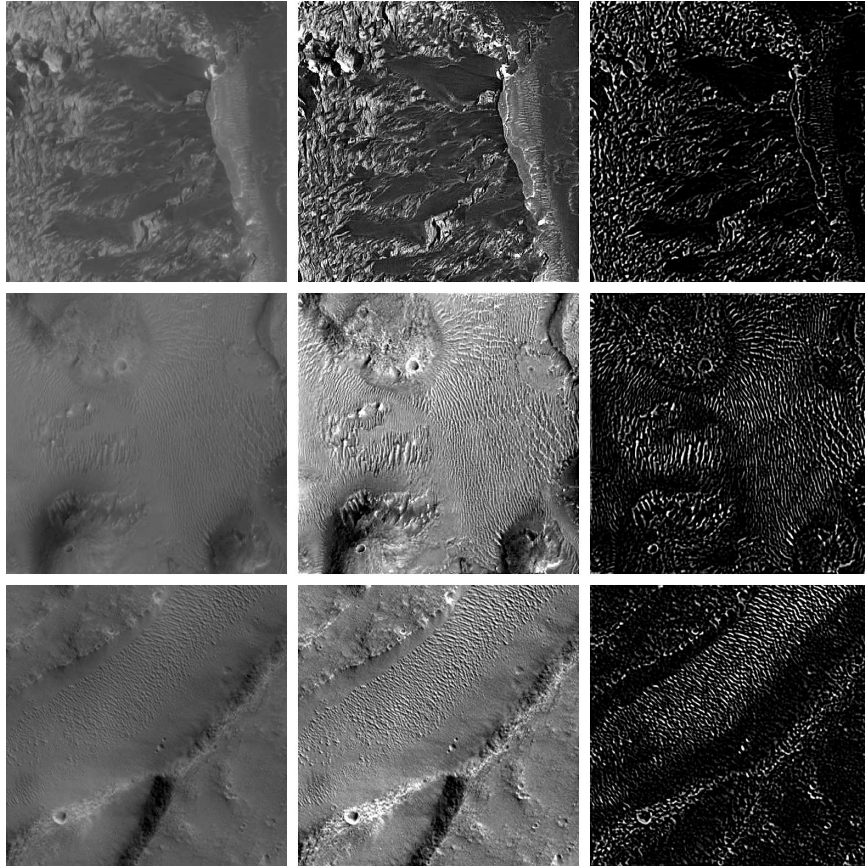
Figure 12. Terrain detection performed on images from Mars Global Surveyor: (left-column) original; (middle-column) SRVR output; (right-column) terrain map.
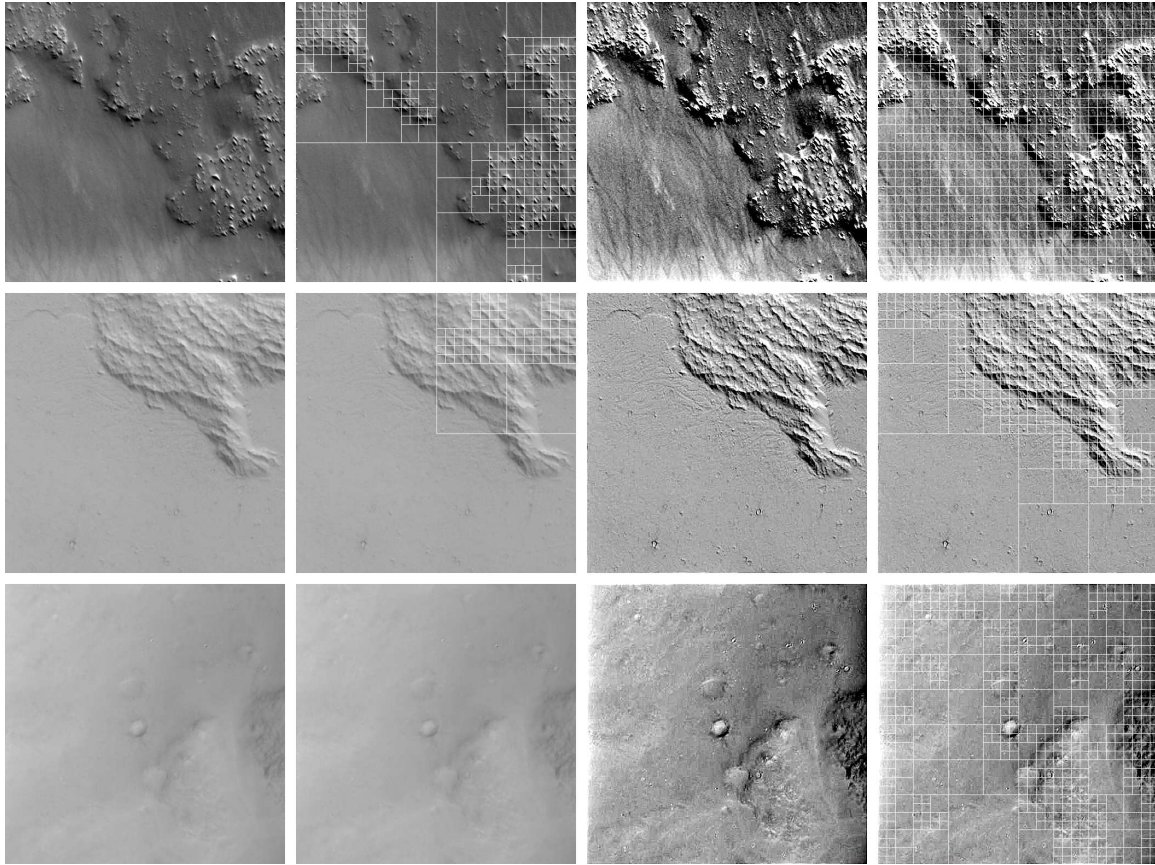
Figure 13. Illustration of the overall concept proceeding to automate landing site selection from a sample of image data from the Mars Global Surveyor imager.

REFERENCES

1. D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multi-scale Retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing: Special Issue on Color Processing* **6**, pp. 965–976, July 1997.

2. Z. Rahman, D. J. Jobson, and G. A. Woodell, "Resliency of the multiscale retinex image enhancement algorithm," in *Proceedings of the IS&T Sixth color Imaging Conference: Color Science, Systems, and Applications*, pp. 129–134, IS&T, 1998.

3. Z. Rahman, D. J. Jobson, and G. A. Woodell, "Retinex processing for automatic image enhancement," *Journal of Electronic Imaging* **13**(1), pp. 100–110, 2004.

4. Z. Rahman, G. A. Woodell, and D. J. Jobson, "A comparison of the multiscale retinex with other image enhancement techniques," in *Proceedings of the IS&T 50th Anniversary Conference*, pp. 426–431, IS&T, 1997.

5. D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. on Image Processing* **6**, pp. 451–462, March 1997.

6. D. J. Jobson, Z. Rahman, and G. A. Woodell, "Feature visibility limit in the nonlinear enhancement of turbid images," in *Visual Information Processing XII*, Z. Rahman, R. A. Schowengerdt, and S. E. Reichenbach, eds., Proc. SPIE 5108, 2003.

7. D. J. Jobson, Z. Rahman, G. A. Woodell, and G. D. Hines, "The automatic assessment and reduction of noise using edge pattern analysis in nonlinear image enhancement," in *Visual Information Processing XIII*, Z. Rahman, R. A. Schowengerdt, and S. E. Reichenbach, eds., Proc. SPIE 5438, 2004.

8. G. A. Woodell, D. J. Jobson, Z. Rahman, and G. D. Hines, "Enhancement of imagery in poor visibility conditions," in *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IV*, E. Carapezza, ed., Proc. SPIE 5778, 2005.

9. Jet Propulsion Laboratories, "Mars global surveyor." http://marsprogram.jpl.nasa.gov/mgs/, 2005.

10. European Space Agency, "Mars express: Multimedia mars express images gallery." http://www.esa.int/SPECIALS/Mars_Express/, 2005.

11. Y. Shostak, Y. Ding, and V. A. Casagrande, "Neurochemical comparison of synaptic arrangements of parvocellular, magnocellular, and koniocellular geniculate pathways in owl monkey (*aotus trivirgatus*) visual cortex," *Journal of Comparative Neurology* **456**, pp. 12–28, 2003.

12. L. P. O'Keefe, J. B. Levitt, D. C. Kiper, R. M. Shapley, and J. A. Movshon, "Functional organization of owl monkey lateral geniculate nucleus and visual cortex," *Journal of Neurophysiology* **80**, pp. 594–609, August 1998.

13. N. H. Yabuta and E. M. Callaway, "Functional streams and local connections of layer4c neurons in primary visual cortex of the macaque monkey," *Journal of Neuroscience* **18**, pp. 9489–9499, November 1998.

14. Z. Rahman, D. J. Jobson, G. A. Woodell, and G. D. Hines, "Impact of multi-scale retinex computation on performance of segmentation algorithms," in *Visual Information Processing XIII*, Z. Rahman, R. A. Schowengerdt, and S. E. Reichenbach, eds., Proc. SPIE 5438, 2004.

15. Y. Itakura, T. Suteo, and T. Takagi, "Statistical properties of the background noise for the atmospheric windows in the intermediate infrared region," in *Infrared Physics*, **14**, pp. 17–29, Pergamon Press, New York, NY, 1974.

16. M. Kass and J. Hughes, "A stochastic model for AI," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 369–372, IEEE, (New York, NY), 1983.

17. F. O. Huck, C. L. Fales, and Z. Rahman, *Visual Communication: An Information Theory Approach*, Kluwer Academic Publishers, Norwell, MA, 1997.